

# CISER

Cornell Institute for Social and Economic Research  
A LEADER IN SOCIAL SCIENCE DATA AND COMPUTING



# Restricted Data Services for Cornell and Beyond



William C. Block, Ph.D.  
Warren A. Brown, Ph.D

IASSIST 2013  
Cologne, Germany  
29 May 2013

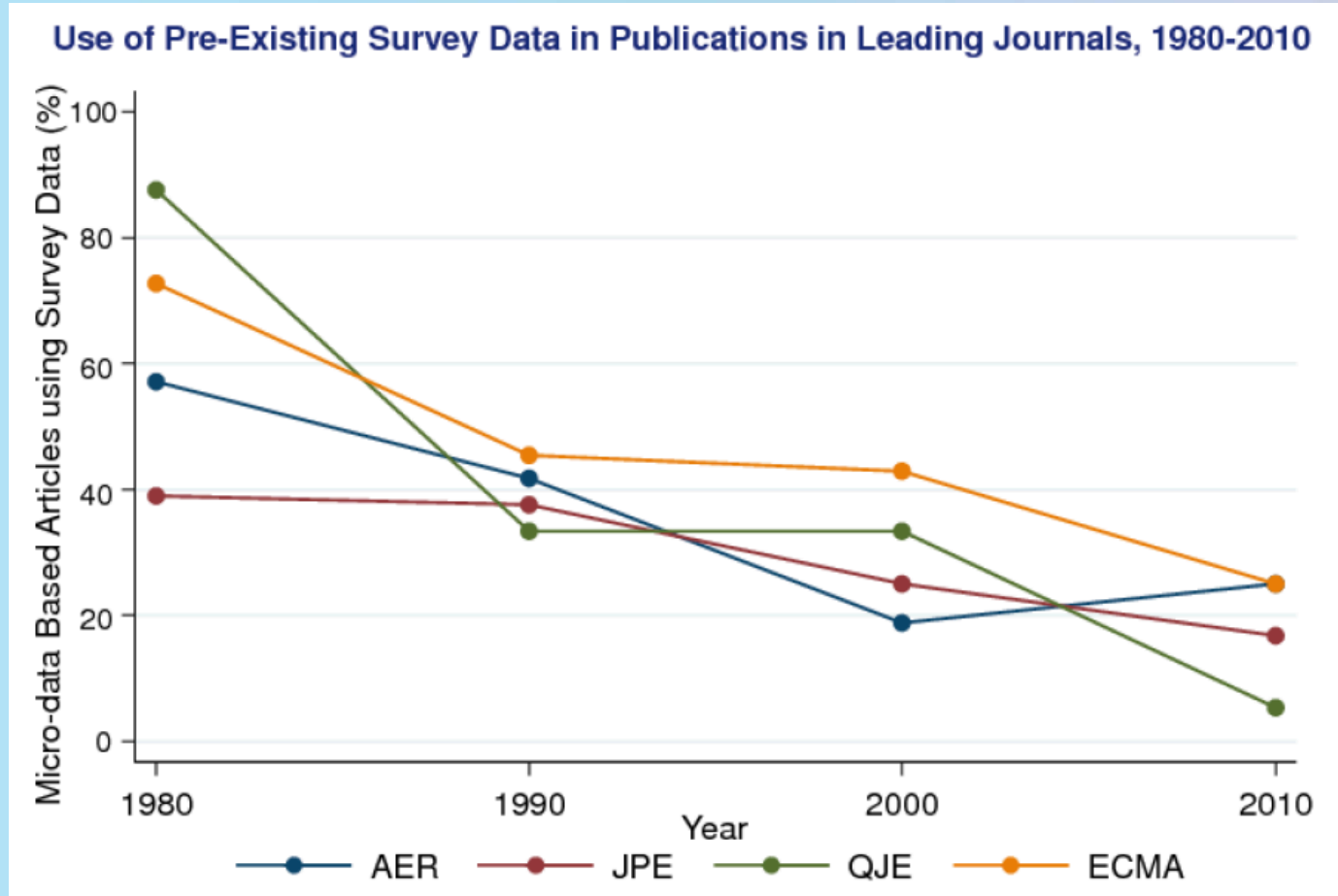


Cornell University

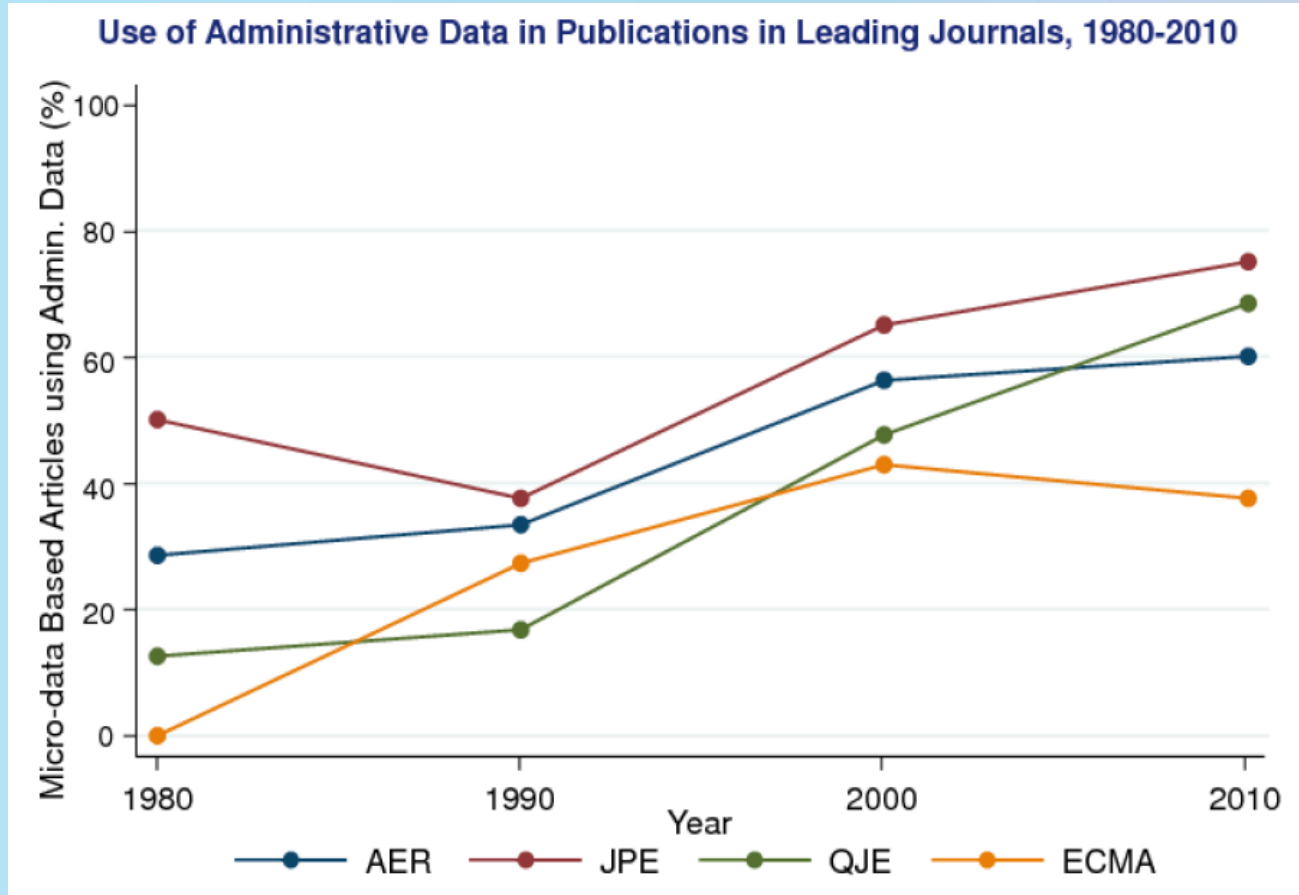
## The Death Knell for Public-use Data

- Sounded by young scholars pursuing research programs that mandate inherently identifiable data: geospatial relations, exact genome data, networks of all sorts, linking administrative records.
- These researchers acquire authorized restricted access to the confidential identifiable data and perform their analyses in secure environments.
- But they don't leave behind the scientific trail that has made public-use files so important.

## Declining use of Public Data in Research



## Increasing use of Restricted Data in Research



# CISER

## Cornell's Official Custodian of Restricted Access Data

Cornell Restricted Access Data Center (CRADC) piloted via an NSF grant in 1999

### CISER's Mission:

....anticipate and support the evolving computational and data needs of Cornell social scientists and economists throughout the entire research process and data life cycle.



## Multiple Modes of Secure Access

- Secure Remote Access
- Secure Rooms/Dedicated Stand-alone computers
- Secure Rooms/Thin-client access to remote-servers
  - Cornell Census Research Data Center (RDC)
  - Institut für Arbeitsmarkt- und Berufsforschung (IAB)
- Cornell Virtual RDC
  - Synthetic Data
  - Scientific integrity, public, training environment

## Computing Security

- Firewall
  - CISCO Advanced Security Appliances (ASA)
  - Federal Information Processing Standard 140-A (FIPS)
- Servers
  - National Institute of Standards and Technology ([NIST](#))
  - Meets DoD C-2 standards for trusted computing environments
- Data
  - Follow [FIPS -199](#) regarding confidential data
  - Health Insurance Portability and Accountability Act of 1996 (HIPAA)



## Computing Hardware

Hardware includes:

- DELL R810 servers
- 40-core
- 512Gb RAM
- 10Gb internal network with failover
- Disk-to-disk backup for secure domain for faster retrieval
- Servers available upon request for specialized projects
- Secured by CISCO Advanced Security Appliances (ASA)



## Software Available on CISER Restricted Access Servers

- 25+ statistical analysis software applications available
  - SPSS, SAS, Stata, R, Matlab, JMP, ArcGIS...
- 17+ other software programs available
  - EndNote, Excel/Word, Python, WinZip, Adobe Professional...
- Ability to access and utilize ~\$100K of software without purchasing or managing updates and licensing

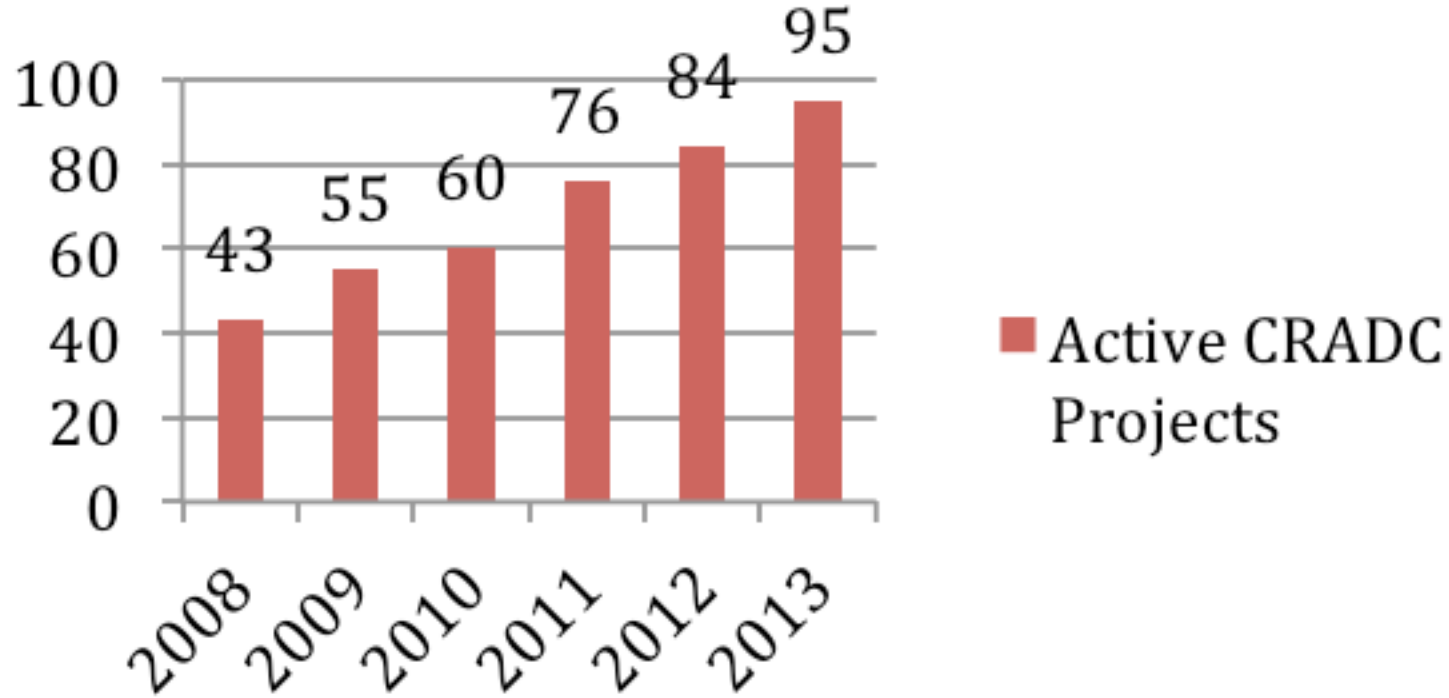
## Secure Data Services Census Research Data Center

- One of 15 Census RDCs in the nation
- Access to confidential microdata.
  - Business data - establishment level, longitudinal
  - Individual and household data- full sample, detailed geography
  - Health data in partnership with NCHS and AHRQ
- Rigorous approval process by Census Bureau
  - Scientific merit
  - Improve data quality
- Able to collaborate with researchers at other RDCs

## Wide Range of Services for Secure Data

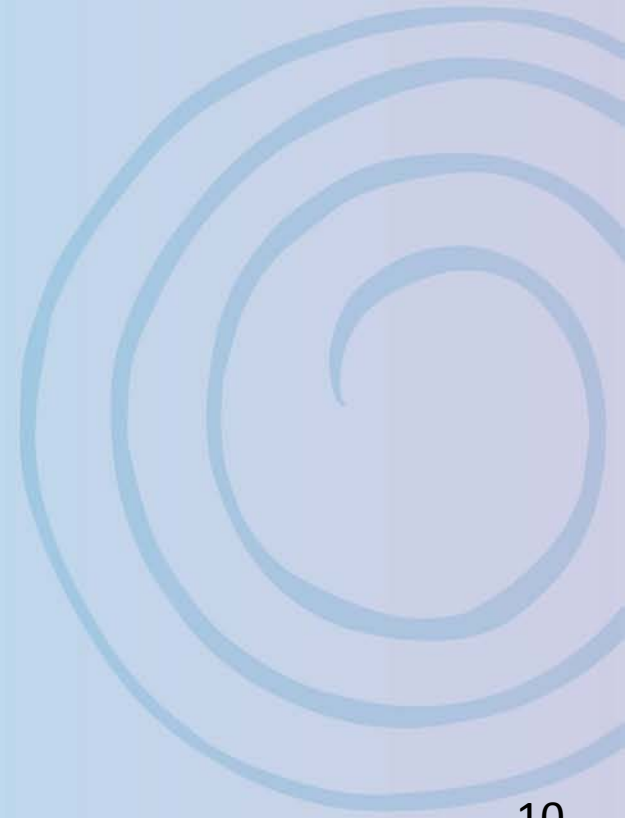
- Identification of applicable data sources
- Writing security plans to meet requirements of data providers
- Coordination with Office of Sponsored Programs (OSP) and Cornell's Institutional Review Board (IRB)
- Design of sponsor-required research data management plans
- Work with data provider for secure transfer of data to CRADC servers
- Disclosure avoidance review where required
- Customized user environment and backup routines
- Audit and site-visit support

## Growth in CRADC Projects



## Seven Characteristics of CISER Secure Data Services

- Comprehensive
- Centralized
- Collaborative
- Competent
- Creative
- Cutting-edge
- Cool



# CISER

## Advancing CISER's Secure Data Services

- **Cornell NSF-Census Research Node**: addressing the tension between restricted data and open scientific inquiry
- **Institute für Arbeitsmarkt- und Berufsforschung (IAB)**: advancing access to secure data across international boundaries (Joerg Heining, 15:45 today, Room Gereon)
- **FISMA**: the Federal Information Security Management Act of 2002
- **Growing External Collaboration**



Thank you for your  
time & attention!

Questions?

[block@cornell.edu](mailto:block@cornell.edu)

